

Fehlerannotierte Lernerkorpora des Deutschen: Methodologie und Empirie

Tagungsprogramm

per Zoom*

Freitag, 22. September 2023

9:00–9:15	Begrüßung
9:15–10:00	Annette Portmann, Christine Renker, Josef Ruppenhofer, Matthias Schwendemann, Katrin Wisniewski und Torsten Zesch: „ Explorative automatisierte Annotationsverfahren syntaktischer Erwerbsstufen des Deutschen im Projekt DAKODA: Ein Werkstattbericht “
10:00–10:45	Christian Horn und Monika Moravkova: „ Das geplante ‚Maus-Korpus‘ – Konzeption, Komponenten und Implementierung “
10:45–11:00	Kaffeepause
11:00–11:45	Dang Thi Thu Hien und Le Tuyet Nga: „ Aufbau des vietnamesischen Lerner*innenkorpus: Konzeption - Datenaufbereitung - Nutzungspotenzial “
11:45–12:30	Anja Binanzer, Hagen Hirschmann und Miriam Langlotz: „ Lernersprachliche Variation bei Satzverknüpfung in schriftlichen Erzählungen ein- und mehrsprachiger Lernender des Deutschen “
12:30–14:00	Mittagspause
14:00–14:45	Krisztina Molnár: „ Lernerkorpora in der Tertiärsprachenforschung “
14:45–15:30	Susan Schlotthauer und Beata Trawiński: „ Propositionale Argumente in deutschen Lernertexten “
15:30–15:45	Kaffeepause
15:45–16:30	Péter Kappel und Bernadett Modrián-Horváth: „ Wortstellungsphänomene in Lernerkorpora “
16:30–17:15	Andreas Nolda: „ Dulko & EXMARaLDA: Von einer externen Toolsammlung zum integrierten Bestandteil “

Samstag, 23. September 2023

9:00–9:15	Begrüßung
9:15–10:00	Yuan Li: „ Aufbau des Chinesischen Deutschlerner-Korpus (CDLK) und die korpusbasierte Erforschung der Lernaltersprache “
10:00–10:45	Nam Luu: „ Analyse der Übersetzungsfehler und deren Ursachen in Prüfungstexten der Abschlussprüfung der Fachübersetzungssseminar vom Jahrgang 2016 der Fakultät für deutsche Sprache und Kultur der ULIS-VNU “
10:45–11:00	Kaffeepause
11:00–11:45	Thi Bao Van Ho: „ Kohäsion und Kohärenz in Abschlussarbeiten vietnamesischer Bachelor- und Masterstudierenden: Eine korpusgestützte Analyse “
11:45–12:30	Helena Wedig, Carola Strobl, Tanja Mortelmans und Jim J. J. Ureel: „ Gebrauch von Kohäsionsmitteln im Deutschen als Fremdsprache von niederländischen Lernenden: Konnektoren und Koreferenz im Beldeko-Korpus “
12:30–14:00	Mittagspause
14:00–14:45	Ewa Drewnowska-Vargáné: „ Gebrauch von ‚Anaphern‘ in den Essays des fehlerannotierten deutsch-ungarischen Dulko-Korpus “
14:45–15:30	Orsolya Rauzs: „ Orthografische Fehler in Dulko- und Falko-Texten “
15:30–15:45	Kaffeepause
15.45–16.30	Abschlussdiskussion

* Zoom-Link auf Anfrage an lernerkorpusstagung2023@gmail.com

Explorative automatisierte Annotationsverfahren syntaktischer Erwerbsstufen des Deutschen im Projekt DAKODA: Ein Werkstattbericht

Annette Portmann (Universität Leipzig)

Christine Renker (Universität Bamberg)

Josef Ruppenhofer (Fernuniversität Hagen)

Matthias Schwendemann (Universität Leipzig)

Katrin Wisniewski (Universität Bamberg)

Torsten Zesch (Fernuniversität Hagen)

Kolleg:innen aus dem Fach DaF/DaZ besitzen häufig nur eingeschränkte Kenntnisse zu Techniken im Umgang mit (großen Mengen an) digitalen Daten. Gerade im Bereich der Lernerkorpuslinguistik nimmt die Bedeutung solcher Kompetenzen jedoch ständig zu. Dieses Desiderat versucht das Projekt DAKODA (*Datenkompetenzen in DaF/DaZ: Exploration sprachtechnologischer Ansätze zur Analyse von L2-Erwerbsstufen in Lernerkorpora des Deutschen*) anzugehen, das in der Richtlinie zur Förderung von Projekten zur *Stärkung der Datenkompetenzen* des wissenschaftlichen Nachwuchses vom BMBF gefördert wird (Laufzeit 10/2022–09/2025). Das interdisziplinäre Vorhaben (Computerlinguistik an der FernUni Hagen, Deutsch als Fremd- und Zweitsprache an der Universität Bamberg) verfolgt das Ziel, Datenmanagement- und Analysekompetenzen zur Konsolidierung großer und heterogener L2-Datenmengen an den Bamberger Nachwuchs zu vermitteln. Außerdem soll ein Grundverständnis sprachtechnologischer Analyseverfahren weitergegeben werden, das die DaF/DaZ-Forschenden mit fundamentalen Schritten zur Entwicklung entsprechender Werkzeuge vertraut macht und in die Lage versetzt, mit diesen Tools selbständig zu arbeiten.

Die Übermittlung von Datenkompetenzen erfolgt dabei im Kontext konkreter Fragestellungen, die den Erwerb syntaktischer Eigenschaften des Deutschen als Fremd- und Zweitsprache betreffen. DAKODA setzt sich hier zum Ziel, die *Nützlichkeit automatisierter Analyseverfahren für die Untersuchung von Erwerbsstufen in Lernersprache zu explorieren*. Vor allem in der Processability Theory (PT; vgl. u. a. Pienemann 1998, 2005; Lenzing et al. 2019) wird der Erwerb zentraler Wortstellungscharakteristika als stufenförmig begriffen. Hinreichende Erklärung für das Erreichen einer Erwerbsstufe ist in der PT je ein einzelner ‚Kernindikator‘. Lernende haben eine Stufe erreicht, wenn sie deren Kernindikator „in principle“ (Pienemann 1998, S. 138) kognitiv verarbeiten können. Diese Erwerbsstufen sind intensiv beforscht (z. B. Haberzettl, 2005, Diehl et al., 2000 und viele andere), äußerst weit verbreitet und kommen auch in didaktischen bzw. diagnostischen Kontexten regelmäßig zum Einsatz. In DAKODA stehen die aufeinander aufbauenden Stufen SEP (Kernindikator: Verbseparation), INV (Subjekt-Verb-Inversion) und VEND (Verbendstellung im Nebensatz)

sowie die Stufe TOP (Topikalisierung von Kernargumenten; Bettoni & Di Biase 2015) im Vordergrund.

Problematisch ist bezüglich der Stufen jedoch zum Beispiel (a) die Tatsache, dass vorliegende Studien oft nur auf Produktionsdaten weniger einzelner Lernender beruhen. Hinzu kommt (b), dass Variation, die in aktuellen Theorieansätzen der L2-Erwerbsforschung (wieder) eine sehr zentrale Rolle spielt, im PT-Framework nur eine geringe bzw. sehr spezifische Bedeutung hat. So werden lernendeninterne und -externe Einflussfaktoren kaum systematisch in ihrer Auswirkung auf Erwerbsverläufe untersucht. Schließlich verstellt (c) die Konzentration auf Kernindikatoren u.E. potenziell den Blick darauf, dass diese unter Umständen in bestimmten linguistischen Kontexten häufiger/seltener/überhaupt (erstmalig) auftreten könnten (z. B. Inversion in bestimmten Vorfeldtypen, Verbseparation mit bestimmten Verben bzw. Verbtypen usw.). Zwar hat Pienemann bereits 1998 darauf verwiesen, dass möglicherweise innerhalb der Erwerbsstufen feinere Erwerbsschritte ausdifferenziert werden könnten; empirische Studien haben sich diesen Fragen für das Deutsche aber kaum gewidmet. Außerdem ist sehr wenig darüber bekannt, inwiefern die Kernindikatoren (d) mit anderen Konzeptualisierungen von „Fortschritt“ im L2-Erwerb zusammenhängen, wie etwa der sprachlichen Komplexität oder Kompetenzniveaus des GER. Inwieweit diese Problemkomplexe in DAKODA angegangen werden können, hängt letztlich vom Erfolg der explorativen automatisierten Analyseprozesse ab, die im Projekt entwickelt werden.

DAKODA bringt derzeit zunächst eine *große Menge an L2-Korpora* zusammen und konsolidiert diese technisch mit dem Ziel, korpusübergreifende Suchen zu ermöglichen. Die Korpusdaten werden je nach datenschutz- und lizenzrechtlichen Gegebenheiten sowohl in einem Repositorium als auch über ein Dashboard zugänglich sein. Das Dashboard soll sich auf Analysefunktionen zu den Erwerbsstufen beschränken.

Zur Datenbasis gehören publizierte Korpora aus dem DaF- (z. B. FALKO-Korpora (Reznicek et al. 2012), Dulko (Beeh et al. 2021), MERLIN (Abel et al. 2014)) und DaZ-Bereich (z. B. ZISA-Daten (Meisel 2020), das ESF-Korpus (Perdue 1993)); einige Korpora enthalten L1-Baselines (z. B. DISKO (Wisniewski et al., 2022; <https://home.uni-leipzig.de/sprastu/korpora/>), MULTILIT (Schroeder et al. 2015)), und auch Herkunftssprecher:innendaten sind in DAKODA enthalten (RUEG-Korpus; Wiese et al. 2021). Andere Datensammlungen sind bisher ganz unveröffentlicht. Es handelt sich insgesamt um eine äußerst vielschichtige und heterogene Datenbasis. Korpusübergreifende Suchen sollen über ein harmonisiertes *Metadatenchema* zugänglich gemacht werden, das wir in unserem Vortrag gern vorstellen würden. Auch auf Herausforderungen aus dem *Preprocessing* (POS-Tagging, Parsing) würden wir gern eingehen.

Sodann wird unser Vortrag auf *erste Schritte bei der explorativen automatischen Analyse* von Erwerbsstufen eingehen. Hierzu wird in DAKODA ein iterativer Workflow mit mehreren qualitativen und quantitativen Validierungsschleifen zum Einsatz kommen, um die Genauigkeit der entwickelten Vorgehensweisen einschätzen zu können. Wir werden uns auf fundamentale Fragen, Herausforderungen und Lösungsvorschläge konzentrieren, die unter anderem die Segmentierung oder die Frage der Definition lernersprachlicher Produktivität betreffen. Dabei wird gezeigt, dass viele vermeintlich praktische Annotationsproblematiken im Kern theoretischer Natur sind.

Literatur

- Bettoni, C. & Biase, B. D. (2015). *Grammatical development in second languages. Exploring the boundaries of Processability Theory*. Benjamins.
- Diehl, E., Christen, H., & Leuenberger, S. (2000). *Grammatikunterricht: Alles für der Katz? Untersuchungen zum Zweitsprachenerwerb Deutsch*. Niemeyer.
- Haberzettl, S. (2005). *Der Erwerb der Verbstellungsregeln in der Zweitsprache Deutsch durch Kinder mit russischer und türkischer Muttersprache*. Narr.
- Lenzing, A. Nicholas, H., & Roos, J. (2019). *Widening contexts for Processability Theory: Theories and issues*. Benjamins.
- Pienemann, M. (1998). *Language processing and second language development. Processability theory*. Benjamins.
- Pienemann, M. (Hrsg.). (2005). *Cross-linguistic aspects of processability theory*. Benjamins.

Korpora

- Abel, A., Wisniewski, K., Nicolas, L., Boyd, A., Hana, J. & Meurers, D. (2014). A Trilingual Learner Corpus illustrating European Reference Levels. *Ricognizioni – Rivista di Lingue, Letterature e Culture Moderne*, 2(1), 111–126. Zugriff am 15.01.2019. <https://doi.org/10.13135/2384-8987/702>
- Beeh, C., Drewnowska-Vargáné, E., Kappel, P., Modrián-Horváth, B., Nolda, A., Rauzs, O., & Scheibl, G. (2021). *Dulko-Handbuch : Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0*. SZTE Germán Filológiai Intézet. <http://acta.bibl.u-szeged.hu/74061/>
- Meisel, J. M. (2020). ZISA. <https://doi.org/10.25592/UHHFDM.1463>
- Perdue, C. (ed.) (1993). *Adult Language Acquisition. Vol 1: Field Methods*. Cambridge University Press. <https://doi.org/10.21415/T55P41>
- Reznicek, M., Lüdeling, A., Krummes, C., Schwantuschke, F., Walter, M., Schmidt, K., Hirschmann, H., Andreas, T. (2012): *Das Falko-Handbuch. Korpusaufbau und Annotationen Version 2.01*, <https://hu-berlin.de/annis-falko>
- Schroeder, C., Schellhardt, C., Akinci, M.-A., Dollnick, M., Dux, G., Gülbeyaz, E. I., Jähnert, A., Koç-Gültürk, C., Kühmstedt, P., Kuhn, F., Mezger, V., Pfaff, C., & Ürkmez, B. S. (2015). *Multilit: Manual, criteria of transcription and analysis for German, Turkish and English*. <https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/8039>
- Wiese, H., Alexiadou, A., Allen, S., Bunk, O., Gagarina, N., Iefremenko, K., Jahns, E., Klotz, M., Krause, T., Labrenz, A., Lüdeling, A., Martynova, M., Neuhaus, K., Pashkova, T., Rizou, V., Rosemarie, T., Schroeder, C., Szucsich, L., Tsehaye, W., Zerbian, S. & Zuban, Y. (2021). *RUEG Corpus*. <https://doi.org/10.5281/ZENODO.3236068>
- Wisniewski, K., Muntchick, E., Portmann, A. (2022). *DISKO (Deutsch im Studium: Lernerkorpus)*. <https://hdl.handle.net/10932/00-0534-6404-3CE0-0001-3>
- Wisniewski, K., Muntchick, E., Portmann, A. (2022). Schreiben in der Studiersprache Deutsch: Das Lernerkorpus DISKO. In Wisniewski, K., Lenhard, W., Möhring, J., Spiegel, L. (eds.). *Sprache und Studienerfolg bei Bildungsausländer/-innen*. Münster: Waxmann.

Das geplante „Maus-Korpus“ – Konzeption, Komponenten und Implementierung

Christian Horn (Hankuk University of Foreign Studies, Seoul)

Monika Moravkova (Hankuk University of Foreign Studies, Seoul)

Der Beitrag erläutert das Konzept zur Erstellung des Korpus des sogenannten „Maus-Projekts“ (vgl. Horn & Moravkova 2018). Bei diesem internationalen, langfristig angelegten Kooperationsprojekt für Deutschstudierende mit der Struktur eines „Plattformprojekts“ (vgl. Horn 2019) spielt die Maus aus der in Deutschland bekannten „Sendung mit der Maus“ die Rolle einer deutschen Touristin, die Deutschlernende in verschiedenen Ländern besucht. Über ihre gemeinsamen Erlebnisse mit der Maus verfassen die Studierenden Texte auf Deutsch und veröffentlichen sie nach der Einarbeitung des Feedbacks ihrer Lehrkraft auf der eigens dafür angelegten Projektwebseite (www.maus-reisen.de). Derzeit sind 32 Lehrkräfte mit etwa 900 Studierenden aus 14 Ländern an dem Projekt beteiligt.

Im Rahmen des Projekts entstehen Texte von Deutschlernenden, die drei Besonderheiten aufweisen: Erstens führen die zahlreichen mitwirkenden Studierenden zu einer Vielzahl von Lernertexten zum Rahmenthema, deren Anzahl sich über den Projektverlauf weiter erhöht. Zweitens sind diese Texte durch die Internationalität des Projekts von Studierenden mit einer Vielzahl unterschiedlicher L1 verfasst. Drittens liegen alle Texte in digitaler Form vor und können so unproblematisch für weitere, auch automatische, Analysen genutzt werden.

Die Ausgangsannahme für die Erstellung des vorgesehenen Korpus besteht darin, dass die Auswertung der im Projekt entstehenden Texte insbesondere wegen der zweiten genannten Besonderheit wertvolle Einsichten in eine Reihe von (fremdspachen)linguistischen Phänomenen liefern würde, etwa hinsichtlich verschiedener Fehlertypen oder Interferenzen zwischen der L1 und der Zielsprache Deutsch. Daher sollen die Lernertexte künftig automatisch erfasst und systematisch in einem ersten Schritt in einer Textsammlung zusammengetragen werden. Dabei sind jedoch mehrere Herausforderungen zu beachten, die im Vortrag thematisiert werden sollen, u. a. die Frage, wie ausschließlich unkorrigierte Texte erfasst werden können, wie dabei von Studierenden verwendete Medien (Fotos/Videos), auf die im Text verwiesen wird, eingebunden werden, und wie das gesamte Verfahren für Studierende, Lehrkräfte und für die Ablage in der Datenbank möglichst niederschwellig und doch effektiv erfolgt. Im Vortrag wird dazu eine Konzeption vorgestellt, die die Grundlage für die Implementierung auf der Projektwebseite darstellt und künftig die automatische Erfassung der Texte sicherstellen soll. Darüber hinaus wird das geplante, begleitende Verfahren zur Erhebung der zugehörigen Metadaten der Studierenden erläutert, das sich vom klassischen Fragebogenformat löst, in den Projektablauf didaktisch integriert ist und so für die Studierenden ein einschlägiges, unauffälliges Element des Projekts darstellen soll. Die Auswahl der abgefragten Lerner-Metadaten orientiert sich dabei an den Vorschlägen in Granger & Paquot (2017), die etwa auch für das Dulko-Korpus (Beeh et al. 2021) verwendet wurden, wobei jedoch insbesondere für die Sprachstandserhebung statt eines standardisierten Tests ein integriertes, komprimiertes Verfahren eingesetzt wird.

In einem späteren zweiten Schritt, der im Vortrag aber nur skizziert wird, soll aus der entstandenen Textsammlung kriterienbasiert ein balanciertes Korpus erstellt werden, das im dritten Schritt auf verschiedenen Ebenen annotiert wird. Dabei ist vorgesehen, dass sich die Annotation prinzipiell an den Richtlinien des Falko-Lernerkorpusprojekts (Reznicek et al. 2012) orientiert. Das annotierte Korpus soll schließlich der wissenschaftlichen Fachgemeinschaft für Analysen zur Verfügung gestellt werden.

Im Vortrag werden schwerpunktmäßig das geplante Vorgehen und die Komponenten zur Erhebung der Daten für die vorgesehene Textsammlung vorgestellt.

Literatur

Beeh, Christoph; Drewnowska-Vargáné, Ewa; Kappel, Péter; Modrián-Horváth, Bernadett; Nolda, Andreas; Rauzs, Orsolya & Scheibl, György (2021): *Dulko-Handbuch. Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0.* Szeged: Institut für Germanistik.

Granger, Sylviane & Magali Paquot (2017): *Core metadata for learner corpora. Draft 1.0.* Manuscript. Louvain-la-Neuve: Université catholique de Louvain.

Horn, Christian (2019): Vom Unterrichtsprojekt zum Plattformprojekt: Potenziale und Erfolgsfaktoren für die Skalierung. *DaF-Szene Korea*, 49, 51–75.

Horn, Christian & Moravkova, Monika (2018): Lach- und Sachgeschichten in Korea – Planung und Reflexion eines Unterrichtsprojekts mit koreanischen Studierenden. *DaF-Szene Korea*, 46, 55–72.

Reznicek, Marc; Lüdeling, Anke; Krummes, Cedric; Schwantuschke, Franziska; Walter, Maik; Schmidt, Karin; Hirschmann, Hagen & Torsten Andreas (2012): *Das Falko-Handbuch. Korpusaufbau und Annotationen Version 2.01.* Berlin: Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik – Korpuslinguistik.

Aufbau des vietnamesischen Lerner*innenkorpus: Konzeption– Datenaufbereitung–Nutzungspotential

Dang, Thi Thu Hien (Universität Hanoi)

Le, Tuyet Nga (Nationaluniversität Hanoi)

Lerner*innenkorpora sind systematische Sammlungen von authentischen (schriftlichen oder mündlichen) digitalisierten Sprachprodukten der Lernenden einer Sprache. Sie bieten eine sehr wichtige Datengrundlage für die empirische Erforschung des Zweit- und Fremdsprachenerwerbs. Für das Fach Deutsch als Fremdsprache sind seit ca. zwei Jahrzehnten einige Lerner*innenkorpora aufgebaut worden, die einen bedeutenden Beitrag zur Erforschung des Erwerbs der deutschen Sprache der Lerner*innen unterschiedlicher Herkünfte geleistet haben (Falko, GeWiss, Merlin usw.). Für die Erforschung des Erwerbs des Deutschen als Fremdsprache der vietnamesischen Lernenden lag jedoch bis 2000 noch keine fundierte Korpus-Sammlung vor. Um diese Lücke zu schließen, wird seit 4 Jahren in zwei Deutschabteilungen in Hanoi an dem Aufbau eines vietnamesischen Lerner*innenkorpus (VIELKO) gearbeitet. Das Korpus besteht aus mündlichen sowie schriftlichen Prüfungstexten und wissenschaftlichen Arbeiten von Studierenden des Studiengangs „Deutsche Sprache“ an der Universität Hanoi und an der Hochschule für Sprachen und Internationale Studien der vietnamesischen Nationaluniversität Hanoi. Im vorliegenden Beitrag wird über den Prozess des Aufbaus dieses Korpus berichtet (Konzeption, Datensammlung, Digitalisierung, Annotation).

Lernersprachliche Variation bei Satzverknüpfung in schriftlichen Erzählungen ein- und mehrsprachiger Lernender des Deutschen

Anja Binanzer (Leibniz-Universität Hannover)

Hagen Hirschmann (Humboldt-Universität zu Berlin)

Miriam Langlotz (Universität Kassel)

Semantische Relationen zwischen Sätzen zu versprachlichen, stellt für Lernende mit Deutsch als Erst- und Zweitsprache im Spracherwerb bzw. der literalen Sprachentwicklung eine komplexe Lernaufgabe dar. Für die sprachliche Realisierung spezifischer Relationen zwischen sprachlichen Äußerungen – z. B. Temporal, Kausal- oder Adversativbeziehungen – steht im Deutschen ein breites Inventar an Junktoren zur Verfügung (vgl. bspw. von Polenz 2008, 252 ff.). Die Beherrschung dieser Funktionswörter bedeutet, für die entsprechenden lexikalischen Einheiten die jeweiligen semantischen Relationen zuordnen zu können, in angemessenen sprachlichen Kontexten einen passenden Junktor zu wählen und gemäß der jeweils geltenden syntaktischen Verwendungsregeln einzubetten (vgl. Binanzer/Langlotz 2019).

Im Projekt JuNaKo (JUnktion–NArration–KOrpuslinguistisch) ist es das Ziel, den schriftsprachlichen Erwerb von Junktoren unter Berücksichtigung genau der erwähnten Verwendungsaspekte zu untersuchen und dabei typische Erwerbsverläufe, individuelle Unterschiede sowie strukturelle Erwerbsschwierigkeiten offenzulegen und zu diskutieren (vgl. Binanzer/Hirschmann/Langlotz 2022). Diese Spezifika stellen auch bei der korpuslinguistischen Annotation eine Herausforderung dar, die auf den verschiedenen Ebenen (lexikalisch, syntaktisch, semantisch) erfolgen muss.

Im geplanten Vortrag stellen wir das Narrations-Lerner-Korpus (NaLeKo) vor, das sich seit 2020 im Aufbau befindet und neben allgemeinen grammatischen und erwerbsspezifischen Analysen korpuslinguistische Ausweisung zur lexikalischen, syntaktischen und semantischen Ausweisung des Gebrauchs von Junktoren beinhaltet. Das Korpus setzt sich aus schriftlichen narrativen, auf Deutsch verfassten Texten (N = 288) von Kinder und Jugendlichen im Alter von sieben bis siebzehn Jahren in den Klassenstufen 2–9. Die Texte stammen von einsprachig mit L1 Deutsch Aufgewachsenen (N = 164), mehrsprachig mit L1 Deutsch Aufgewachsenen (N = 55) und mit Deutsch als L2 Aufgewachsenen (69). Die Texte wurden zu Bildimpulsen bzw. spezifischen Schreibaufgaben elizitiert. Ausgehend von oben dargestellten Forschungsanliegen stellen wir erste Ergebnisse zu der Frage dar, wie die Lernenden temporale Satzverknüpfungen versprachlichen.

Im letzten Teil des Vortrags gehen wir aus methodischer Perspektive auf lernersprachliche Varianten als Untersuchungsgegenstand selbst und als Herausforderung für die Korpusannotation ein. „Fehler“ im Bereich der Junktions sind kein trivialer Untersuchungsgegenstand, da bei der Versprachlichung ein gewisses Spektrum von Ausdrucksvarianten und verschiedenen syntaktischen Konstruktionen besteht, so dass

Bezüge zwischen lernersprachlichen Varianten und bestimmten zielsprachlichen Varianten nicht immer klar auf der Hand liegen, was die Interpretation der lernersprachlichen Äußerung erschwert.

Literatur

Binanzer, A. & Langlotz, M. (2019): Junktion und Narration. Schreibentwicklungsprozesse ein- und mehrsprachiger Kinder. In A. Binanzer, M. Langlotz & V. Wecker (Hrsg.): Grammatik in Erzählungen – Grammatik für Erzählungen. Erwerbs-, Entwicklungs- und Förderperspektiven. Baltmannsweiler: Schneider, 125–150.

Binanzer, A., Hirschmann, H. & Langlotz, M. (2022): Narrative Funktionen temporaler Junktoren – Lernertexte und Kinder- und Jugendliteratur im korpuslinguistischen Vergleich. In: Mesch, B.; Uhl, B. (Hrsg.): Temporalität – empirische Zugänge zum Erwerb von Zeitlichkeit. Münster: Waxmann, 175–198.

Von Polenz, P. (2008): Deutsche Satzsemantik. Grundbegriffe des Zwischen-den-Zeilen-Lesens. Berlin, New York: De Gruyter.

Lernerkorpora in der Tertiärsprachenforschung

Krisztina Molnár (Universität Pécs)

Im Vortrag wird der Frage nachgegangen, ob und inwieweit fehlerannotierte Lernerkorpora in der Tertiärsprachenforschung verwendet werden können.

Der Terminus „Tertiärsprache“ wird in der einschlägigen Literatur auf verschiedene Weisen definiert, in der vorliegenden Arbeit wird jede Fremdsprache als Tertiärsprache bezeichnet, die nach der ersten Fremdsprache gelernt wird.

Die Tertiärsprachenforschung ist aus der Erkenntnis entstanden, dass eine zweite Fremdsprache (und auch alle weiteren Fremdsprachen) anders erlernt wird als die erste Fremdsprache. Eine wichtige Fragestellung ist, wie die bereits bekannten Sprachen das Erlernen einer neuen Sprache beeinflussen. Bei verschiedenen Untersuchungen wurde jedoch auch festgestellt, dass auch die neuen Sprachen Einfluss auf die schon bekannten Sprachen haben können (vgl. Jarvis & Pavlenko 2008: 20).

Im Vortrag werden Ergebnisse einer Pilotanalyse gezeigt. Dabei werden Belege aus den beiden Subkorpora von DULKO (vgl. Beeh et al. 2021) analysiert, die von ungarischen Deutschlernenden stammen, die als weitere Fremdsprache Englisch sprechen. Es wird einerseits untersucht, ob Interferenzfehler vorhanden sind, deren mögliche Ursache das Englische ist. Andererseits werden die beiden Subkorpora miteinander verglichen, da es anzunehmen ist, dass Essays und Übersetzungen eine unterschiedliche Anzahl an Interferenzfehler enthalten.

Literatur

Beeh, Christoph, Drewnowska-Vargáné, Ewa, Kappel, Péter, Modrián-Horváth, Bernadett, Nolda, Andreas, Rauzs, Orsolya & Scheibl, György (2021): Dulko-Handbuch. Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0. Szeged: Institut für Germanistik.

Jarvis, Scott & Pavlenko, Aneta (2008): Crosslinguistic influence in language and cognition. New York/London: Routledge.

Propositionale Argumente in deutschen Lernertexten

Susan Schlotthauer (Leibniz-Institut für Deutsche Sprache, Mannheim)

Beata Trawiński (Leibniz-Institut für Deutsche Sprache, Mannheim)

Mit unserem Beitrag versuchen wir, einen Brückenschlag zwischen der Lernerkorpusforschung und der sprachvergleichenden Forschung herzustellen, und gehen der Frage nach, welche Reflexe übereinzelsprachlicher Varianz sich in Lernertexten finden. Die Studie steht im Kontext des Projekts „Grammatik des Deutschen im europäischen Vergleich – Verbgrammatik“ am Leibniz-Institut für Deutsche Sprache, das propositionale Argumente im Deutschen aus sprachvergleichener Perspektive untersucht. Als Kontrastsprachen dienen dabei das Englische, Italienische, Polnische und Ungarische. Unsere Datengrundlage sind Lernertexte des Deutschen, die den Korpora Dulko (Essay- und Übersetzungstexte), Falko und Merlin (jeweils nur Essaytexte) entnommen wurden. In die Betrachtung einbezogen wurden alle Texte, die von Muttersprachlern des Englischen, Italienischen, Polnischen und Ungarischen verfasst worden.

Zwei Themenkomplexe werden behandelt:

- die Varianz zwischen finiten und infiniten Formen in Objektfunktion (für das Deutsche, Italienische und Ungarische beschrieben u. A. von Brandt/Bildhauer 2019; K. Molnár 2016, 2018; Wöllstein 2015)
 - a. zum einen bei Strukturen mit Referenzidentität zwischen den Subjekten des Matrixprädikats und des eingebetteten Prädikats:
 - (1) *Ich habe endlich so entschieden, dass ich 10 Tage in der Hauptstadt verbringen werde.* (Korpus: Dulko, Übersetzung, Muttersprache: Ungarisch)
 - (2) *Deshalb viele Menschen eint sich einen Einfamilienhaus zu bauen in die kleine Orte neben der Stadt.* (Korpus: Merlin, Essay, Muttersprache: Polnisch)
 - b. zum anderen bei Strukturen ohne Referenzidentität zwischen den Subjekten des Matrixprädikats und des eingebetteten Prädikats:
 - (3) *Endlich haben sie mir erlaubt, allein nach einem so weit liegenden Land zu fahren.* (Korpus: Dulko, Übersetzung, Muttersprache: Ungarisch)
 - (4) *Ich würde für jede Leser empfehlen neue Kulturen kennenzulernen.* (Korpus: Merlin, Essay, Muttersprache: Ungarisch)
- das (Nicht-)Vorkommen von propositionalen Proformen (Korrelaten) zu Objekt- und Subjektsätzen (in Bezug auf die Einzelsprachen und teilweise auch kontrastiv beschrieben u. A. von Mollica 2010; V. Molnár 2015; Paranhos Zitterbart 2002; Sudhoff 2003, 2016)

- (5) *Ich mag auch natürlich mit anderen Gebräuchen kennen lernen, aber ich erwarte das, dass das fremde Land über mein Land, und über unsere Tradition wenige Dinge kennt.* (Korpus: Merlin, Essay, Muttersprache: Ungarisch)
- (6) *Wenn es ihm gesagt wurde, er solle zum Kühlschrank gehen ...* (Korpus: Dulko, Übersetzung, Muttersprache: Ungarisch)

Für beide Phänomenbereiche können drei Einflussfaktoren nachgewiesen werden, die die Wahl der Realisierungsform des propositionalen Arguments und die (Nicht-)Setzung des Korrelats in den Lernertexten steuern:

- die Muttersprache des Lerner: Texte ungarischer Muttersprachler unterscheiden sich in beiden Punkten deutlich von den Texten von Lernern anderer Muttersprache,
- der Typus des Lernertexts: Essay- und Übersetzungstexte (ungarischer Muttersprachler) unterscheiden sich deutlich,
- die allgemeine Vorkommensfrequenz der Matrixprädikate.

Gezeigt wird, wie diese Faktoren interagieren und wie sich die Muttersprache der Lerner nicht nur in Fehlern, sondern auch in den gewählten Versprachlichungsstrategien und -präferenzen widerspiegelt. Somit ist der Beitrag auch als Plädoyer für den Einsatz von Lernerkorpora in der sprachvergleichenden Forschung zu verstehen.

Literatur

- Brandt, Patrick & Bildhauer, Felix (2019): Alternation von *zu-* und *dass-*Komplementen: Kontrolle, Korpus und Grammatik. In: Fuß, Eric, Konopka, Marek & Wöllstein, Angelika (Hg.): *Grammatik im Korpus. Korpuslinguistisch-statistische Analysen morphosyntaktischer Variationsphänomene*. Tübingen: Narr (Studien zur deutschen Grammatik, 80), 211–297.
- Mollica, Fabio (2010): *Korrelate im Deutschen und im Italienischen*. Frankfurt am Main [u.a.]: Lang (Deutsche Sprachwissenschaft international, 9).
- Molnár, Krisztina (2016): Infinitive und Infinitivkonstruktionen im Deutschen, Italienischen und Ungarischen. In: *Sprachwissenschaft* 41 (2), 185–212.
- Molnár, Krisztina (2018): *Infinitivkonstruktionen als Objekte im Sprachvergleich*. Berlin, Bern & Wien: Lang (Szegediner Schriften zur germanistischen Linguistik, 8).
- Molnár, Valéria (2015): The Predicationality Hypothesis. The case of Hungarian and German. In: É. Kiss, Katalin, Surányi, Balázs & Dékány, Éva (Hg.): *Papers from the 2013 Piliscsaba conference*. Amsterdam: Benjamins (Approaches to Hungarian, 14), 209–245.
- Paranhos Zitterbart, Jussara (2002): *Zur korrelativen Subordination im Deutschen*. Berlin/Boston: De Gruyter (Linguistische Arbeiten, 464).
- Sudhoff, Stefan (2003): *Argumentsätze und es-Korrelate. Zur syntaktischen Struktur von Nebensatzeinbettungen im Deutschen*. Berlin: wvb.
- Sudhoff, Stefan (2016): Correlates of object clauses in German and Dutch. In: Frey, Werner, Meinunger, André & Schwabe, Kerstin (Hg.): *Inner-sentential propositional proforms. Syntactic properties and interpretative effects*. Amsterdam/Philadelphia: Benjamins (Linguistik aktuell, 232), 23–48.
- Wöllstein, Angelika (2015): Grammatik – explorativ. Hypothesengeleitete und -generierende Exploration variierender Satzkomplementationsmuster im standardnahen

Deutsch. In: Eichinger, Ludwig (Hg.): *Sprachwissenschaft im Fokus. Positionsbestimmungen und Perspektiven*. Berlin, Boston: De Gruyter (Jahrbuch Institut für Deutsche Sprache, 2014), 93–120.

Wortstellungsphänomene in deutschen Lernerkorpora

Bernadett Modrián-Horváth (Universität Szeged)

Péter Kappel (Universität Szeged)

“The L1 tunes the ways in which learners attend to language. Learned-attention transfers to L2 and it is this L1 entrenchment that limits the end-state of usage-based SLA. But these limitations can be overcome [...]” (Ellis 2019: 47)

In der Fremdsprachenerwerbsforschung kommt der Wortstellung in der Zielsprache Deutsch eine besondere Aufmerksamkeit zu, denn sie gilt aus mehrerer Hinsicht als sprachtypologisch „exotisch“ und stellt somit eine Herausforderung für Deutschlernende dar (vgl. Riehl 2018: 254). Unser Vortrag geht der Frage nach, inwieweit muttersprachliche Wortstellungsmuster den Erwerb der deutschen Wortstellung bei Sprachlernenden unterschiedlicher Muttersprachen (fortan L1) beeinflussen. Im Vortrag soll thematisiert werden, welche syntaktischen Normen auf einem hohen Sprachniveau (B2 und darüber hinaus) noch verletzt werden und wie häufig diese Normverstöße bei den einzelnen Probandengruppen auftreten.

Die Analyse stützt sich auf eine empirische Untersuchung im *Deutsch-ungarischen Lernerkorpus* (Dulko) sowie in anderen vergleichbaren Korpora (Falko, Kobalt-DaF und MERLIN). Der Vergleich erfolgt jeweils mit Daten aus Teilkorpora, die aufgrund der L1 der Probanden definiert wurden. Aus sprachtypologischer Perspektive interessant sind hier z. B. das Schwedische, das wie das Deutsche über eine V2-Stellung verfügt, und das Englische, in dem die Subjekterststellung weit überwiegt, während Ungarisch eine Sprache mit informationsstrukturell motivierter Wortstellung ist (Molnár 2012). Unter die Lupe genommen werden lernersprachliche Phänomene im Bereich der Verbstellung (in Haupt- und Nebensätzen). Zudem werden weitere Wortstellungsphänomene stichprobenartig behandelt.

Literatur

Beeh, Christoph / Drewnowska-Vargáné, Ewa / Kappel, Péter / Modrián-Horváth, Bernadett, Nolda, Andreas / Rauzs, Orsolya / Scheibl, György (2021): Dulko-Handbuch. Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0. Szeged: Institut für Germanistik. DOI: [10.14232/dulko-handbuch-v1.0](https://doi.org/10.14232/dulko-handbuch-v1.0)

Boyd, Adriane / Hana, Jirka / Nicolas, Lionel / Meurers, Detmar / Wisniewski, Katrin / Abel, Andrea / Schöne, Karin / Štindlová, Barbora / Vettori, Chiara. (2014). The MERLIN corpus: Learner language and the CEFR. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, S. 1281–1288, Reykjavik, Iceland. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2014/pdf/606_Paper.pdf

Ellis, Nick C. (2019): Essentials of a Theory of Language Cognition. In: *The Modern Language Journal*, 103 (Supplement 2019). DOI: [10.1111/modl.12532](https://doi.org/10.1111/modl.12532)

Reznicek, Marc / Lüdeling, Anke / Krummes, Cedric / Schwantuschke, Franziska / Walter, Maik / Schmidt, Karin / Hirschmann, Hagen / Andreas, Torsten (2012): *Das Falko-Handbuch. Korpusaufbau und Annotationen Version 2.01*. Berlin: Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik – Korpuslinguistik.

[https://www.linguistik.hu-](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2/)

[berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2/](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2/)

(Stand: 1. 6. 2022)

MERLIN – Multilingual Platform for European Reference Levels: Interlanguage Exploration in Context (Technische Universität, Dresden).

Molnár, Valéria (2012): Zur Relevanz der linken Peripherie für die Strukturierung der Information – kontrastive und typologische Überlegungen. In: Gunkel, Lutz/Zifonun, Gisela (Hrsg.): *Deutsch im Sprachvergleich. Grammatische Kontraste und Konvergenzen*. Berlin, Boston: de Gruyter, S. 383–416. (Jahrbuch des Instituts für Deutsche Sprache 2011), DOI:

[10.1515/9783110284768.383](https://doi.org/10.1515/9783110284768.383)

Riehl, Claudia Maria (2018): Deutsch aus kontrastiver Perspektive. In: Harr, Anne-Katharina (Hrsg.): *Deutsch als Zweitsprache*. Stuttgart: J. B. Metzler, S. 237–274.

Zinsmeister, Heike / Reznicek, Marc / Ricart Brede, Julia / Rosén, Christina / Skiba, Dirk (2012): Das Wissenschaftliche Netzwerk "Kobalt-DaF". In: *Zeitschrift für Germanistische Linguistik* 40 (3), S. 457–458. DOI: [10.1515/zgl-2012-0030](https://doi.org/10.1515/zgl-2012-0030)

Dulko & EXMARaLDA: Von einer externen Toolsammlung zum integrierten Bestandteil

Andreas Nolda (Berlin-Brandenburgische Akademie der Wissenschaften)

Das „[Deutsch-ungarische Lernerkorpus \(Dulko\)](#)“ (Beeh *et al.* 2021) ist ein Korpus mit deutschsprachigen Lernertexten ungarischer Muttersprachler, das im Rahmen des Dulko-Projekts am Institut für Germanistik der Universität Szeged erstellt und 2021 auf dem [ANNIS-Server](#) des Lehrstuhls für Korpuslinguistik und Morphologie an der Humboldt-Universität zu Berlin für die Recherche verfügbar gemacht wurde. Das dafür entwickelte Annotationsverfahren von Hirschmann und Nolda (2019) baut auf dem Annotationsverfahren der Falko-Korpora (Reznicek *et al.* 2012) auf, geht aber im Hinblick auf die Fehlerannotation wesentlich darüber hinaus. So ist das Dulko-Korpus nicht nur mit Wortarten, Lemmata, Satzspannen, Zielhypothesen und Abweichungen annotiert, sondern auch durchgängig mit Fehlerkategorien aus einem mehrdimensionalen Fehler-Tagset. Außerdem können beliebig viele Zielhypothesen angegeben werden, um sich überlappende Fehler zu repräsentieren (Nolda 2019, 2023).

Für die Annotation des Dulko-Korpus wurde der EXMARaLDA-Partitureditor (Schmidt *et al.* 2016) eingesetzt unter Verwendung einer Toolsammlung namens „[EXMARaLDA \(Dulko\)](#)“. Mit Hilfe der dort mittels XSLT-Stylesheets implementierten Transformationsszenarien können Lernertexte automatisch tokenisiert und mit Wortarten, Lemmata und Satzspannen annotiert werden. Außerdem werden Annotatoren bei der manuellen Angabe von Zielhypothesen und Fehlerkategorien unterstützt; die Berechnung von Abweichungen zwischen Lernertext und Zielhypothesen erfolgt wiederum automatisch. Für Lemmatisierung und Wortarten-Tagging wird der TreeTagger (Schmid 1997) aufgerufen. EXMARaLDA (Dulko) wurde 2018 mit dem Innovationspreis der Universität Szeged ausgezeichnet und trotz der komplexen Installation und Konfiguration von mehreren anderen Lernerkorpusprojekten nachgenutzt.

2020 wurden die zentralen Komponenten von EXMARaLDA (Dulko) in das [EXMARaLDA-GitHub-Repositoryum](#) integriert; sie werden Teil der nächsten offiziellen EXMARaLDA-Version sein. Damit erübrigt sich eine separate Installation von EXMARaLDA (Dulko). Die Konfiguration der TreeTagger-Anbindung wurde wesentlich vereinfacht und erlaubt die Nutzung des TreeTaggers für andere Sprachen als das Deutsche. Dazu kommen diverse Verbesserungen der linguistischen Datenverarbeitung und der Nutzerführung.

Dieser Vortrag zeigt, wie Lernerdaten in der neuen EXMARaLDA-Version mit den Dulko-Transformationsszenarien annotiert werden können und welche neuen Möglichkeiten sich dabei bieten.

Literatur

Beeh, Christoph *et al.* (2021). Dulko-Handbuch: Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0. Manuskript, Universität Szeged. <http://acta.bibl.u-szeged.hu/74061/> [25. Nov. 2021].

- Hirschmann, Hagen und Andreas Nolda (2019). Dulko – auf dem Weg zu einem deutsch-ungarischen Lernerkorpus. In *Neues vom heutigen Deutsch: Empirisch – methodisch – theoretisch*, hg. v. Ludwig Eichinger und Albrecht Plewnia, Institut für Deutsche Sprache: Jahrbuch 2018, Berlin: de Gruyter, 339–342.
- Nolda, Andreas (2019). Annotation von Lernerdaten mit EXMARaLDA (Dulko). Manuskript, Berlin: Berlin-Brandenburgische Akademie der Wissenschaften. https://andreas.nolda.org/publications/nolda_2019_annotation_lernerdaten.pdf [26. Feb. 2019].
- Nolda, Andreas (2023). Fehlerannotation und Fehleranalyse am Beispiel des deutsch-ungarischen Lernerkorpus Dulko. In *Jahrbuch für internationale Germanistik: Beihefte*, Bd. 10, 747–755. <https://www.peterlang.com/document/1277913> [28. Feb. 2023].
- Reznicek, Marc *et al.* (2012). Das Falko-Handbuch: Korpusaufbau und Annotationen. Version 2.01. Manuskript, Humboldt-Universität zu Berlin. <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2> [26. Jan. 2017].
- Schmid, Helmut (1997). Probabilistic part-of-speech tagging using decision trees. In *New Methods in Language Processing*, hg. v. Daniel B. Jones und Harold L. Somers, London: Routledge, 154–164.
- Schmidt, Thomas *et al.* (2016). EXMARaLDA Partitur-Editor: Manual. Version 1.6. Manuskript. http://www.exmaralda.org/pdf/Partitur-Editor_Manual.pdf [27. März 2018].

Aufbau des Chinesischen Deutschlerner-Korpus (CDLK) und die korpusbasierte Erforschung der Lernaltersprache

Yuan Li (Zhejiang Universität)

Die deutsche Sprache erlebt in China seit der Jahrtausendwende einen Boom. Der Erhebung 2020 zufolge gibt es 20 % mehr Deutschlerner als noch vor 5 Jahren. Die bestehenden Forschungsprojekte zur Lernaltersprache der chinesischen Deutschlernenden verfolgten meist einen qualitativen und induktiven Ansatz und stützten sich selten auf umfangreiche Korpora. Weltweit gesehen enthalten nur einige Lernerkorpora schriftliche Texte der chinesischen DeutschlernerInnen. Es ist zu erkennen, dass es bei den aktuellen Korpora nur begrenzte Textsorten und nicht umfassende Deutschniveaus sowie Lernstufen gibt.

Angesichts dieses Desiderats wird ein umfangreiches „Chinesische Deutschlerner-Korpus“ (CDLK) an der Zhejiang Universität aufgebaut. Das ist ein multidimensional annotiertes Lernerkorpus, das alle Niveaus bzw. Lernstufen abdeckt und die Texte sowohl von SchülerInnen, als auch von Studierenden enthält. Darüber hinaus werden die Texte von denselben Lernenden innerhalb von drei Jahren regelmäßig gesammelt.

Auf der Grundlage des CDLK wurden bereits einige Studien durchgeführt. Auf der lexikalischen Ebene wurden die Verwendung von Konnektoren, Verben auf Basis der Valenzgrammatik, Präpositionen, Chunk sowie lexikalische Fehler untersucht. Auf der syntaktischen Ebene konzentrieren wir uns derzeit auf die Entwicklung der syntaktischen Komplexität, die Wortstellung der Verben und die syntaktischen Vielfalt. Die Transfers von der Muttersprache und Englisch als erste Fremdsprache auf Deutsch und die retroactiven Transfers vom Deutschen auf Englisch werden ebenfalls thematisiert.

Literatur

Granger, Sylviane (2004): Computer learner corpus research: Current status and future prospects. In: *Language and Computers*, S. 123–145.

Leech, Geoffrey (1997): *Teaching and Language Corpora: A Convergence*. In: Wichmann, Anne, Fligelstone, Steven, McEnery, Tony & Knowles, Gerry (Hg.): *Teaching and language Corpora*. London: Routledge.

Li, Yuan & Ge, Nannan (2019): Digitalisierung in der chinesischen Germanistik am Beispiel der Korpuslinguistik im digitalen Zeitalter In: *Jahrbuch für Internationale Germanistik*. Heft 2.

Li, Yuan et al. (2021): *Das Handbuch CDLK. Version 2.0.* (unveröffentlichtes Manuskript).

Lüdeling, Anke & Kytö, Merja (Hg.) (2009): *Corpus Linguistics. An International Handbook*. Vol. 2. (Reihe Handbücher zur Sprach- und Kommunikationswissenschaft). Berlin: Mouton de Gruyter.

Nolda, Andreas (2019): Annotation von Lernerdaten mit EXMARaLDA (Dulko). Berlin: Berlin-Brandenburgische Akademie der Wissenschaften. https://andreas.nolda.org/publications/nolda_2019_annotation_lernerdaten.pdf (Stand: 20. 4. 2022)

Reznicek, Marc, Lüdeling, Anke, Krummes, Cedric, Schwantuschke, Franziska, Walter, Maik, Karin, Schmidt, Hirschmann, Hagen & Andreas, Torsten (2012): Das Falko-Handbuch. Korpusaufbau und Annotationen Version 2.01. <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/falko/FalkoHandbuchV2/> (Stand: 20. 4. 2022)

Reznicek, Marc, Lüdeling, Anke & Hirschmann, Hagen (2013): Competing target hypotheses in the Falko corpus: A flexible multi-layer corpus architecture. In: Díaz-Negrillo, Ana, Ballier, Nicolas & Thompson, Paul (Hg.): Automatic Treatment and Analysis of Learner Corpus Data. Amsterdam/Philadelphia: John Benjamins, S. 101–124.

Schiller, Anne, Teufel, Simone, Stöckert, Christine & Thielen, Christine (1999): Guidelines für das Tagging deutscher Textcorpora mit STTS (kleines und großes Tagset). Stuttgart: Universität Stuttgart, Institut für maschinelle Sprachverarbeitung und Universität Tübingen, Seminar für Sprachwissenschaft. <http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf> (Stand: 20. 4. 2022)

Wu, Zekun & Li, Yuan (2022a): Zur syntaktischen Komplexität des Schriftdeutschen chinesischer Deutschlerner(innen) – Eine korpusbasierte Profilanalyse. In: Deutsch als Fremdsprache, Heft 4, S. 207–217.

Wu, Zekun & Li, Yuan (2022b): Der Gebrauch von Konnektoren bei chinesischen DaF-Lernern. Eine korpuslinguistische Untersuchung argumentativer Lernertexte. In: Studien zur deutschen Sprache und Literatur. Heft 2, S. 111–140.

Analyse der Übersetzungsfehler und deren Ursachen in Prüfungstexten der Abschlussprüfung der Fachübersetzungsseminar vom Jahrgang 2016 der Fakultät für deutsche Sprache und Kultur der ULIS-VNU

Luu Trong Nam (University of Languages and International Studies, Vietnam National University)

An der Fakultät für Deutsche Sprache und Kultur der Fremdsprachenhochschule der Nationaluniversität Hanoi gilt die Fachrichtung Translation als eine der wichtigsten fachlichen Hauptrichtungen, die von den Studierenden der Fakultät ausgewählt werden kann. Als Dozierende dieser Fachrichtung suchen die Kollegen der Translationsabteilung stets nach Methoden und Didaktisierungsmaßnahmen zur Verbesserung der übersetzerischen Leistung der in diesen Fachbereich konzentrierten Studentinnen und Studenten. In diesem Zusammenhang bietet eine Fehleranalyse der Übersetzungen der Studierenden weitreichende Erkenntnisse über die Hintergründe der bestehenden Fehler, denn wie es in einem deutschen Sprichwort heißt, „aus Fehlern wird man klug“.

Diese Arbeit konzentriert sich daher auf die Fehler, die in den Prüfungstexten des Fachübersetzungsseminars des Jahrgangs 2016 vorkommen und aus der VieLKO-Lernerdatenbank entnommen werden können. Fehler können oftmals im Kontext der Didaktik untersucht werden. Insbesondere im Bereich des Fremdsprachenunterrichts sind sie von Bedeutung, da sie die objektivsten Rückmeldungen der Lernenden darstellen. Diese Fehler ermöglichen es, die Effektivität der Lehrmethoden und -techniken der Lehrkraft zu bewerten und notwendige Anpassungen vorzunehmen, um die Lernziele zu erreichen. Im Bereich der Übersetzungswissenschaft haben zahlreiche Forscherinnen und Forscher Übersetzungsfehler auf Grundlage verschiedener Aspekte analysiert. Nord (2011) beispielsweise, stützte ihre Kategorisierung von Übersetzungsfehlern auf die von ihr definierten Übersetzungsprobleme: pragmatische, kulturpaarspezifische und sprachpaarspezifische Übersetzungsfehler. H. Lee-Jahnke (zitiert nach Delisle et al., 1999) hingegen betrachtete Präzision, Kreativität und Skopos (Ziel der Übersetzung) als Schlüsselaspekte zur Bewertung von Übersetzungen und zur Identifizierung von Fehlern. Stolze (vgl. Beitrag in „Translationsdidaktik, Grundfragen der Übersetzungswissenschaft“, 1997: 592-602) unterteilte Übersetzungsfehler in Klarheits-, Genauigkeits- und Wirkungsfehler. Diese Kategorisierung basiert auf den Kernprinzipien für den praktizierenden

Übersetzer, nämlich Klarheit, Genauigkeit und die Gesamtwirkung der Übersetzung (Lê Hoài n, 2010, zitiert nach Stolze: 92).

Die obengenannten Prüfungstexte sind Klausurübersetzungen von 34 Studierenden im sechsten Semester des Fachbereichs Translation an der Fakultät für deutsche Sprache und Kultur der Fremdsprachenhochschule der VNU wurden als Korpus für diese Untersuchung herangezogen. In dieser Klausur wurde die fachliche Übersetzungskompetenz der Studierenden in Bezug auf die bilinguale Übersetzung (Vietnamesisch-Deutsch und umgekehrt) geprüft und bewertet. Da sich die Übersetzungsrichtung von Vietnamesisch nach Deutsch oftmals als herausfordernd für die Studierenden erweist, wurde diese für die empirische Untersuchung ausgewählt. Die Studierenden starteten auf dem Sprachniveau B2, da sie bereits vier Semester Sprachunterricht erhalten haben. Im fünften Semester beschäftigten sie sich intensiv mit der deutschen Sprache durch die Teilnahme an linguistischen Seminaren, die Themen wie Phonologie, Morphologie und Syntax abdeckten. Zudem wurden in allen vier Semestern (von Deutsch 1A bis 4C) zahlreiche grammatische und thematische Aspekte der deutschen Sprache behandelt. Anschließend sollen anhand der Ergebnisse der Fehleranalyse die möglichen Gründe mit Hilfe der Theorie der Übersetzungsprobleme von Stolze (2008) festgestellt werden. Darauf aufbauend können übersetzungsmethodische Vorschläge empfohlen werden, indem passende Übersetzungsstrategien für das jeweilige Übersetzungsproblem, welches der Grund für die Übersetzungsfehler ist, vorgestellt und vertieft werden.

Keywords: Übersetzungsfehler, Korpuslinguistik, Fehleranalyse, Translation, Übersetzungsproblem

Literaturverzeichnis:

Delisle, J.; Lee-Jahnke, H.; Cormier, M. C.; Albrecht, J. (1999): *Terminologie de la traduction*: J. Benjamins.

Lê Hoài Ân (2010): *Probleme und Lösungsstrategien beim Übersetzen aus dem Vietnamesischen ins Deutsche*. (Dissertation). Hamburg: Universität Hamburg.

Nord, C. (2011): Funktionsgerechtigkeit und Loyalität: *Theorie, Methode und Didaktik des funktionalen Übersetzens*: Frank & Timme (Funktionsgerechtigkeit und Loyalität).

Stolze, R. (2008): *Übersetzungstheorien. Eine Einführung*. München: Narr.

Kohäsion und Kohärenz in Abschlussarbeiten vietnamesischer Bachelor- und Masterstudierenden: Eine korpusgestützte Analyse

Thi Bao Van Ho (Universität Leipzig / University of Languages and International Studies, Vietnam National University)

Das akademische Schreiben umfasst alle Schreibprodukte, die innerhalb eines Studiums produziert werden, und bereitet den unerfahrenen Schreibenden vielerlei Schwierigkeiten beim Schreibprozess, die sich u. a. durch die Annäherung von studentischen Sprachwerken an der konzeptionellen Mündlichkeit charakterisieren lassen. Bei nicht-muttersprachlichen Studierenden (L2-Studierenden) taucht in den meisten Fällen daneben noch eine Vielzahl von Problemen in der textuellen Mikrostruktur auf, was wiederum den Erwerb der wissenschaftlichen Schreibkompetenz beeinträchtigt. Ausgegangen von der Hypothese, dass die Beherrschung der wissenschaftlichen Schreibkompetenz positiv mit der wissenschaftliche Schreiberfahrung der Schreibenden korreliert, lässt sich in meiner Dissertation die Fragen behandelt, welche Ausdifferenzierungen beim wissenschaftlichen Schreiben zwischen L2-Studierenden in Bachelor- und Masterstudiengängen zu betrachten sind und inwiefern sich diese Unterschiede in Bezug auf zwei Indikatoren von Kohäsion und Kohärenz – nämlich Junktion (nach Ágel / Diegelmann 2010) und Satzstruktur bei komplexen Sätze (nach Linke et al. 2004) – in der Gestaltung ihrer Abschlussarbeiten widerspiegeln. Die Daten stammen aus *VieLko* – einem Lernerkorpus mit mündlichen sowie schriftlichen Daten von Germanistik-Studierenden in Vietnam – und werden mithilfe des Korpustools *EXMARaLDA (Dulko)* (Hirschmann / Nolda 2019) gemäß der Auswahl an Indikatoren auf mehreren Ebenen annotiert und anhand der korpusgestützten Methode in *R* analysiert. Im Vergleich zu anderen Arten von Lernertexten bereitet die Textsorte ‚Akademisches Schreiben‘ eine Vielzahl neuer Herausforderungen bei der Aufbereitung der Primärdaten, deshalb wird der Fokus in meinem Beitrag auf die Datenannotation gelegt. Daneben wird auch eine Übersicht über den Projekt-Workflow, über die Datenverwaltung sowie über die Datenanalyse dargestellt.

Keywords: Schreibentwicklung, akademisches Schreiben, Lernerkorpus, korpusgestützte Analyse, Kohäsion, Kohärenz

Ausgewählte Bibliografie

Ágel, Vilmos / Diegelmann, Carmen (2010): Theorie und Praxis der expliziten Junktion. In: Vilmos Ágel (Hg.): *Nähe und Distanz im Kontext variationslinguistischer Forschung*. Berlin: de Gruyter (Linguistik - Impulse & Tendenzen, 35), S. 345–396.

Averintseva-Klisch, Maria (2018): *Textkohärenz*. 2., aktualisierte Auflage. Heidelberg: Universitätsverlag Winter (Kurze Einführungen in die germanistische Linguistik - KEGLI, Band 14).

Hirschmann, Hagen / Nolda, Andreas (2019): Dulko – auf dem Weg zu einem deutsch-ungarischen Lernerkorpus. In: Eichinger, Ludwig / Plewnia, Albrecht (Hrsg.), *Neues vom heutigen Deutsch: Empirisch – methodisch – theoretisch*. Institut für Deutsche Sprache: Jahrbuch 2018, S. 339-342. Berlin: de Gruyter.

Linke, Angelika et al. (2004): *Studienbuch Linguistik*. Ergänzt um ein Kapitel "Phonetik, Phonologie" von Urs Willi. 5., erweiterte Auflage. Tübingen: Max Niemeyer Verlag (Studienbuch, 121).

Portmann, Paul R. (2018): Textstruktur und Textkompetenz. Der Weg Studierender zum wissenschaftlichen Text. In: Sabine Schmölder-Eibinger et al. (Hg.): *Wissenschaftliches Schreiben lehren und lernen. Diagnose und Förderung wissenschaftlicher Textkompetenz in Schule und Universität*. Münster, New York: Waxmann, S. 174–192.

Steinhoff, Torsten (2007): *Wissenschaftliche Textkompetenz. Sprachgebrauch und Schreibentwicklung in wissenschaftlichen Texten von Studenten und Experten*. Tübingen: Niemeyer (Reihe Germanistische Linguistik, 280).

Gebrauch von Kohäsionsmitteln im Deutschen als Fremdsprache von niederländischen Lernenden: Konnektoren und Koreferenz im Beldeko-Korpus

Carola Strobl (University of Antwerp)

Helena Wedig (University of Antwerp)

Tanja Mortelmans (University of Antwerp)

Jim J. J. Ureel (University of Antwerp)

Während Textkohärenz und -kohäsion direkten Einfluss auf die Verständlichkeit von Texten sowie die Bewertung dieser durch ExpertenInnen nehmen (s. Crossley & McNamara, 2010; Wedig & Strobl, 2022), fallen sie auch fortgeschrittenen Lernenden noch deutlich schwer (Crossley, Kyle & McNamara 2016; Lindström & Lubińska, 2022; Takač et al., 2020). So zeigen Studien, dass Lernende zum Beispiel durch ihre Erstsprache (L1) oder andere Fremd- oder Zweitsprachen in der Nutzung von Kohäsionsmitteln in der Fremdsprache (L2) beeinflusst werden können (Breindl, 2016; Hinkel, 2001). Der spezifisch lernersprachliche Gebrauch von Kohäsionsmitteln kann durch manuelle Annotation und anschließende Analyse untersucht werden. Dies ist das Ziel der vorliegenden Studie, die auf dem spezifischen Gebrauch von Konnektoren und Koreferenz von Deutschlernenden mit L1 Niederländisch fokussiert.

Die Studie wurde anhand des bislang einzigen DaF-Korpus von Lernenden mit L1 Niederländisch durchgeführt: Beldeko (*Belgisches Deutschkorpus*, Strobl, 2020). Das Korpus enthält 301 Zusammenfassungen von belgischen DaF-Studierenden. Für das Forschungsprojekt wurde das Korpus mithilfe der Annotationsplattform Inception (Klie et al., 2018) auf der Basis von selbst erstellten Richtlinien mit Annotationen zu Konnektoren (nach PDTB-3: Webber et al., 2019) und zu Koreferenz (nach Becher, 2011; Kunz et al., 2021) angereichert. Ebenso wurden von den drei muttersprachlichen AnnotatorInnen Abweichungen vom muttersprachlichen Gebrauch markiert.

Die vorläufige Analyse der Annotationen via R zeigt, dass die Lernenden Konnektoren häufig in semantisch abweichenden Kontexten verwenden. Die Nutzung bestimmter abweichender Konnektoren („Letztens“, „Zum Schluss“) ist dabei möglicherweise auf die abweichende Semantik ähnlicher Formen im Niederländischen („ten laatste“, „tenslotte“) oder im Englischen („last“) zurückzuführen. Auf dem Gebiet der Koreferenz fallen vor allem ein uneigentlicher Gebrauch von bestimmten Artikeln bei Eigennamen und fehlende Kongruenz bei Pronomina auf. In unserer Präsentation gehen wir näher auf die Fragen ein, (1) auf welchen Ebenen der Kohäsion lernersprachenspezifischer Gebrauch zu finden ist, (2) welche Unterschiede zum muttersprachlichen Gebrauch sich in der Syntax und Semantik zeigen und (3) ob bestimmte Aspekte dabei auf Interferenz aus dem Niederländischen hinweisen.

Literatur

- Becher, V. (2011). When and why do translators add connectives?: A corpus-based study. *Target: International Journal of Translation Studies*, 23(1), 26–47.
- Breindl, E. (2016). Konnexion in argumentativen Texten. Gebrauchsunterschiede in Deutsch als L2 vs. Deutsch als L1. In *Normalität in der Sprache*, 37–64.
- Crossley, S. A., Kyle, K., & Mcnamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1–16.
- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied language learning*, 12(2), 111–132.
- Klie, J.-C., Bugert, M., Boullosa, B., Eckart de Castilho, R. and Gurevych, I. (2018). The INCEption Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation. In Proceedings of System Demonstrations of the 27th International Conference on Computational Linguistics (COLING 2018), Santa Fe, New Mexico, USA.
- Kunz, K., Lapshinova-Koltunski, E., Martínez, J. M. M., Menzel, K., & Steiner, E. (2021). GECCo-German-English Contrasts in Cohesion. De Gruyter Mouton.
- Lindström, E., & Lubińska, D. (2022). Target-like and non-target-like conjunctive relations in L2 Swedish beginner writing. *Linguistics and Education*, 71, 101073.
- Strobl, C. (2020). *Beldeko Summary Corpus v1.0.0*, Eurac Research CLARIN Centre, <http://hdl.handle.net/20.500.12124/15>.
- Takač, V. P., Kružić, B., & Ivezić, S. V. (2020). A Corpus-Driven Exploration of Discourse Markers in L2 Academic Texts. In *Corpus Analysis in Different Genres* (S. 169–190). Routledge.
- Webber, B., Prasad, R., Lee, A., and Joshi, A. (2019). The Penn Discourse Treebank 3.0 Annotation Manual.
- Wedig, H., & Strobl, C. (2022). Die Bewertung von Kohärenz und Kohäsion in narrativen DaF-Texten: Eine korpusbasierte Untersuchung sprachlicher Einflussfaktoren. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 27(1).

Gebrauch von ‚Anaphern‘ in den Essays des fehlerannotierten deutsch-ungarischen Dulko-Korpus

Ewa Drewnowska-Vargáné (ehm. Universität Szeged)

‚Anaphern‘ tragen bekanntlich zur Stiftung des semantischen und konzeptuellen Textzusammenhangs bei. Als textuelles Phänomen werden sie aber im Unterricht Deutsch als Fremdsprache (DaF) kaum bzw. nicht systematisch behandelt. Übrigens dürfte dies nicht nur im Unterricht DaF in Ungarn ein Problem bedeuten (vgl. dazu bspw. Thielmann (2019)).

Aus meinen Erfahrungen und aus den Umfragen unter Kolleginnen und Kollegen folgt, dass ‚Anaphern‘ auch den Studierenden zu Beginn ihres Germanistikstudiums in Ungarn kaum bekannt sind.

Die Funktion von ‚Anaphern‘ bei der Stiftung des Textverstehens wird erst ab dem dritten Semester oder sogar später – meistens in den Veranstaltungen zur Textlinguistik – systematischer behandelt. Dies gilt wohl auch für Studierende am Institut für Germanistik in Szeged mit Ungarisch als Muttersprache und Deutsch als Fremdsprache, die mindestens das Sprachniveau B2 im Deutschen gemäß GER/CEFR erreicht haben. Gemeint sind hier die Autorinnen und Autoren von Essays und Übersetzungen im Dulko-Korpus (vgl. Korpus (2021) und Handbuch (2021)). Den Gegenstand meiner Untersuchung bildet das Essay-Teilkorpus. Die zentralen Fragestellungen bei der Untersuchung richten sich darauf

- (1) welche typischen Fehler im Anapherngebrauch in den Essays des Dulko-Korpus auftreten und inwiefern diese Fehler das Textverstehen beeinträchtigen sowie
- (2) welche von den Fehlern, die in dem Korpus annotiert worden sind, mit dem Gebrauch von ‚Anaphern‘ im Zusammenhang stehen? (Hierbei sei wohlgemerkt, dass ‚Anaphern‘ nicht im Fokus der Fehlerannotation im Dulko-Korpus stehen. Somit sind die Fehlertags keineswegs speziell auf ‚Anaphern‘ ausgerichtet (vgl. dazu Handbuch 2021, Kap. 5: Annotation)).

‚Anaphern‘ fasse ich textfunktional als Ausdrücke auf, die vorwiegend satzübergreifend zur Wiederaufnahme anderer, meistens ihnen im Text vorausgehender Ausdrücke (‚Antezedentien‘) dienen (vgl. dazu v.a. Consten/Schwarz-Friesel 2007; Schwarz-Friesel/Consten 2014; Brinker/Cölfen/Pappert 2014). ‚Vorwiegend Satzübergreifend‘ bedeutet, dass im Zentrum der Untersuchung ‚Normalanaphora‘ stehen – im Unterschied zu ‚Mikroanaphora‘, die im selben Satz fungieren (vgl. Canisius 1997).¹ Falls in den untersuchten Essays auch kataphorisch, d.h. vorausweisend eingesetzte ‚Anaphern‘ vorkommen (vgl. Brinker/Cölfen/Pappert 2014: 55; Schwarz-Friesel/Consten 2014: 121), werden sie ebenfalls berücksichtigt.

¹ ‚Mikroanaphora‘ werden von der Untersuchung nicht ausgeschlossen, weil ihre Analyse ggf. zu forschungsrelevanten Ergebnissen führen kann. Vgl. das Beispiel aus Kleists Michael Kohlhaas in Canisius (1997: 84): „An den Ufern der Havel lebte um die um die Mitte des sechzehnten Jahrhunderts ein Roßhändler, Namens Michael Kohlhaas, Sohn eines Schulmeisters, einer der rechtschaffensten zugleich und entsetzlichsten Menschen seiner Zeit.“ [Hervorhebungen von mir – E.D.V-]. Beispiele für problematische und daher für die Untersuchung interessante ‚Mikroanaphora‘ im Korpus: Entlohnung 5: Satzspanne 19, 21; Entlohnung 6: Satzspanne 10

Das Inventar von Pronomina, von nominalen Phrasen, in bestimmten Fällen aber auch von Adverbien, mit denen ‚Anaphern‘ ausgedrückt werden können, stellt sich recht umfangreich dar. Dafür seien nur einige Beispiele eingeführt. Neben Pronomina der 3. Person (vgl. weiter unten (1) (3) und (6)) umfasst das Inventar auch Nominalphrasen (vgl. (1) (4) und (5)), daneben aber auch das Pronomen *es* in einer abgeleiteten Funktion (vgl. (2)), Demonstrativpronomina (vgl. (3) und (5)), demonstrative Nominalphrasen (vgl. (4) und (5)), Adverbien (vgl. (6)), Pronominaladverbien und Possessivpronomina (vgl. (7)) (im besonderen Hinblick auf ‚Anaphernresolution‘ als Beitrag zum Textverstehen, vgl. Zifonun/Drewnowska 2011):

- (1) Ein Mann geht in zur Polizei. *Er / Der Mann* ist ziemlich aufgeregt.
- (2) Im Zimmer spielten ein Mädchen und ein Junge. *Es* waren seine Kinder.
- (3) Maria begrüßte zunächst Katrin, und dann umarmte *diese sie*.
- (4) In der Nachbarschaft wohnt auch ein Hund. *Dieses arme Tier* ist sehr oft allein zuhause.
- (5) Vor einer Woche hat mich eine Katze gebissen. *Dies / Diesen Vorfall* vergesse ich nie.
- (6) Peter wohnt in Budapest. *Dort* kann *er*, sooft *er* will, in die Oper gehen.
- (7) Ein Kind kann viel erzählen. *Dabei* spielt *seine* Phantasie eine erhebliche Rolle.

Bei der Untersuchung befaße ich mich mit ‚Anaphern‘ auch im Hinblick auf den Umfang ihres Bezugs. So z. B. stellen das Pronominaladverb *dabei* in (7), ferner das Demonstrativum *dies* und die demonstrative Nominalphrase *diesen Vorfall* in (5) ‚komplexe Anaphern‘ dar, da sie sich im Unterschied zu anderen obigen Beispielen jeweils auf den ganzen vorangehenden Satz beziehen.

Schließlich spielt bei der Untersuchung eine Unterscheidung zwischen ‚Identitäts-Anaphern‘ (alle obigen Beispiele) und ‚Kontiguitäts-Anaphern‘ (vgl. z. B. Canisius 1997; Consten /Schwarz-Friesel 2007) ebenfalls eine wichtige Rolle.

Korpus: Beeh, Christoph / Drewnowska-Vargáné, Ewa / Kappel, Péter / Modrián-Horváth, Bernadett / Nolda, Andreas / Rauzs, Orsolya / Scheibl, György (2021): Dulko: Deutsch-ungarisches Lernerkorpus. Version 1.0. Szeged: Universität Szeged, Institut für Germanistik. <https://korpling.org/annis/> [31. 03. 2023].

Handbuch: Beeh, Christoph / Drewnowska-Vargáné, Ewa / Kappel, Péter / Modrián-Horváth, Bernadett, Nolda, Andreas / Rauzs, Orsolya / Scheibl, György (2021): Dulko-Handbuch. Aufbau und Annotationsverfahren des deutsch-ungarischen Lernerkorpus. Version 1.0. Szeged: Institut für Germanistik. DOI: [10.14232/dulko-handbuch-v1.0](https://doi.org/10.14232/dulko-handbuch-v1.0) [31. 03. 2023].

Sekundärliteratur:

Brinker, Klaus/Cölfen, Hermann/Pappert, Steffen (2014): Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. 8. neu bearbeitete und erweiterte Aufl., Berlin: Schmidt.

Consten, Manfred/Schwarz-Friesel, Monika (2007): Anapher. In: Hoffmann, Ludger (Hrsg): Handbuch der deutschen Wortarten. Berlin etc.: de Gruyter, 265–292.

Canisius, Peter (1997): Kp. III:6: Koreferenz und Kontiguität und dreierlei Anaphora. In: Canisius, Peter/Knipf, Elisabeth (1997): Textgrammatik. Ein Einführungskurs. Budapest: Nemzeti Tankönyvkiadó, 80–91.

Drewnowska-Vargáné, Ewa/Zifonun, Gisela (2011): Formensystem und textuelle Verwendung von Possessiva: Ein deutsch-, polnisch- und ungarischsprachiger Paralleltextvergleich. In: Deutsche Sprache 3/11, 207–233 - Berlin: Schmidt. (Deutsche Sprache Jahrgang 3/39).

Schwarz-Friesel, Monika/Consten, Manfred (2014): Einführung in die Textlinguistik. Darmstadt: WBG.

Thielmann, Winfried (2019): Deixis und Anapher in der Sprachdidaktik. In: Zielsprache Deutsch 46, 3(2019), 7–25.

Orthografische Fehler in Dulko- und Falko-Texten

Orsolya Rauzs (Universität Szeged)

Im Vortrag wird der Frage nachgegangen, was für orthografische Fehler in den deutschsprachigen Lernertexten von ungarischen Studierenden mit Deutschkenntnissen auf B2 typischerweise vorkommen. Untersucht werden die Übersetzungen und Essays des Dulko-Korpus (insg. 21.010 Tokens), das an der Universität Szeged (Ungarn) erstellt wurde. Zur orthografischen Fehlerebene gehören im Dulko-Projekt Fehler, die folgende Bereiche betreffen (s. Beeh / Drewnowska-Vargáné / Kappel / Modrián-Horváth / Nolda / Rauzs / Scheibl 2021, S. 33–35):

- Groß- und Kleinschreibung
- Getrennt- und Zusammenschreibung
- Wortschreibung (fehlende oder überflüssige Buchstaben, Verwechslung von Buchstaben)
- Zeichensetzung (falsche Setzung von Satzzeichen auf Satz- und Wortebene)

Diese vier Fehlertypen werden im Vortrag anhand der Dulko-Texte einerseits statistisch erfasst, um festzustellen, womit die untersuchte Studierendengruppe tendenziell die größten Probleme hatte. Andererseits wird durch Beispiele gezeigt, was für spezifische Fehlermuster identifiziert werden können.

Die Ergebnisse werden in einem zweiten Schritt mit ausgewählten Texten des Falko-Teilkorpus falkoEssayL1v2.3 kontrastiert, das Essays von deutschen Gymnasiasten enthält. Dieser Vergleich sollte einen Einblick darin gewähren, welche die wesentlichen Unterschiede zwischen der orthografischen Kompetenz einer bestimmten Gruppe von Muttersprachlern und der ungarischer Deutschlernenden sind.

Literatur

Beeh, Christoph / Drewnowska-Vargáné, Ewa / Kappel, Péter / Modrián-Horváth, Bernadett / Nolda, Andreas / Rauzs, Orsolya / Scheibl, György (2021): Dulko-Handbuch. Version 1.0. Szeged: Institut für Germanistik (http://acta.bibl.u-szeged.hu/74061/1/dulko_handbuch_2021.pdf)

Verwendete Korpora

DulkoEssay-v1.0 (https://korpling.org/annis/#_c=RHVsa29Fc3NheS12MS4w)

DulkoTranslation-v1.0

(https://korpling.org/annis/#_c=RHVsa29UcmFuc2xhdGlvbi12MS4w)

falkoEssayL1v2.3 (https://korpling.org/annis/#_c=ZmFsa29Fc3NheUwxdjluMw)